

This is the final accepted manuscript of:

C. Puzzarini, V. Barone. Diving for Accurate Structures in the Ocean of Molecular Systems with the Help of Spectroscopy and Quantum Chemistry. *Acc. Chem. Res.* 51, 548–556 (2018). DOI: 10.1021/acs.accounts.7b00603

Available at: <https://doi.org/10.1021/acs.accounts.7b00603>

© 2018 American Chemical Society

Diving for accurate structures in the ocean of molecular systems with the help of spectroscopy and quantum chemistry

Cristina Puzzarini^{*,†} and Vincenzo Barone^{*,‡}

*Dipartimento di Chimica “Giacomo Ciamician”, Università di Bologna, Via F. Selmi 2, 40126
Bologna, Italy, and Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

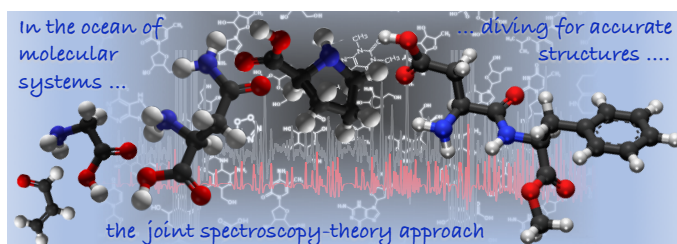
E-mail: cristina.puzzarini@unibo.it; vincenzo.barone@sns.it

^{*}To whom correspondence should be addressed

[†]Dipartimento di Chimica “Giacomo Ciamician”, Università di Bologna, Via F. Selmi 2, 40126 Bologna,
Italy

[‡]Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

CONSPECTUS: The prediction and interpretation of structural properties are the starting points for a deep understanding of thermochemistry, kinetics and spectroscopic sig-



natures of molecular systems. To give an example, a detailed knowledge of the conformational behavior of the main building blocks of biomolecules in the gas phase (i.e., without the perturbing effect of the environment) is a mandatory prerequisite toward the understanding of the role played by different interactions in determining the biological activity in terms of structure-activity relationships. The first step to take is an unambiguous definition of molecular structure. We address the so-called Born-Oppenheimer equilibrium structure, which is defined in a rigorous manner and isotopically independent, and the target accuracy. For the latter, we aim at the so-called “spectroscopic” accuracy, which implies uncertainties of a few milliÅ for bond lengths and smaller than a tenth of degree for angles.

If on one side the continuous enhancements of the experimental techniques give access to new and unprecedented spectroscopic determinations, on the other side they require increasing efforts for an unbiased interpretation and analysis. Among the pieces of information, accurate molecular structures play a particularly important role. Indeed, there is a strong relationship between the experimental outcome and the electronic structure of the system. Spectroscopic techniques, in particular those exploited in the gas phase, are therefore accurate and reliable sources for structural information. However, it is seldom straightforward to derive molecular structures directly from the experimental information. Indeed, even in the favorable case of investigations in the gas phase, vibrational effects are always present and disentangling their contributions requires to collect information for all vibrational modes, a nearly impossible protocols strategies can be identified, which are referred to as “top-down” and “bottom-up”. The first approach, denoted as *semi-experimental approach*, relies on extracting from experimental outcomes the equilibrium structure by using quantum-chemical computations to recover vibrational effects. The “bottom-up” approach consists in verifying the computed equilibrium geometry by means of a comparison between calculated and experimental spectroscopic parameters that probe structural characteristics.

In this contribution, we try to review the most important challenges in accurate molecular structure determinations, with particular emphasis on the “solution” provided by a joint theoretical-experimental approach and on the current state of the art. Starting from the illustration of different strategies, we proceed by addressing the increasing complexity in the derivation of equilibrium geometries: from the construction of a database of accurate structures, we then face the problem of extending the dimension of the systems amenable to accurate structural determinations, and finally we move to the challenge of understanding the nature of intermolecular interactions.

Introduction

Reliable prediction of molecular structures is at the heart of chemistry and spectroscopy, yet remains a challenging task that over the years has attracted the attention of many experimentalists and theoreticians.^{1,2} The first step to take is an unambiguous definition of molecular structure, which is actually provided by quantum mechanics. The so-called Born-Oppenheimer (BO) approximation, decoupling the electrons and nuclei motions, allows the definition of potential energy surfaces (PES), whose minima correspond to equilibrium structures (r_e). Indeed, this definition has a number of advantages: it has well defined ranges of application, excludes vibrational effects in a rigorous manner, and is independent of the considered isotopic species. From another point of view, molecular structures in condensed phases are determined by the interplay of stereo-electronic, dynamic, and environmental effects. It is of paramount importance to disentangle the role of the different contributions in order to tune molecular properties for different technological and/or biological applications.^{3,4}

The focus of this contribution will be the prediction of equilibrium structures for isolated molecules, which is a very demanding task for both pure experimental and theoretical approaches. Focusing on molecular spectroscopy, there is a strong relationship between the experimental outcome and the electronic structure of the system under investigation. Therefore, spectroscopic techniques, in particular those exploited in the gas phase, are accurate sources for structural information.²⁻⁵ However, it is seldom straightforward to derive molecular structures from the experimental information. Even in the most favorable case of investigations in the gas phase, vibrational effects are always present and disentangling their contributions requires to collect information for all vibrational modes, a nearly impossible task for all but the smallest molecular systems.^{1,2,6} For first-principles computations, the challenge is the accuracy because it requires sophisticated quantum-chemical (QC) models together with a full account of core-valence (CV) correlation contributions as well as the extrapolation to the complete basis set (CBS) limit (see, for example,⁷⁻¹⁰ and references therein). Indeed, for medium-to-large sized molecular systems, such computations are not affordable because of the unfavourable scaling of highly correlated levels of theory with the number of basis functions. Although recent works have shown that basis-set and core-valence correlation effects can be effectively evaluated at lower computational levels,^{9,11-13} the dimension

of the systems amenable to these accurate approaches remains quite limited. Under such circumstances, a viable solution is provided by integrated experimental-theoretical strategies, which pave the route toward the extension of accurate structural studies to systems larger than those treatable by experimental and QC methods separately.

We can conceive two joint theory-spectroscopy strategies, which are referred to as “top-down” and “bottom-up”. The first approach, denoted as *semi-experimental approach*,^{2,14,15} relies on extracting from experimental outcomes the equilibrium structure details by using QC computations for providing the missing information. If gas-phase spectroscopic investigations are considered, QC calculations are employed to recover vibrational effects and derive equilibrium structures from experimental data, thus leading to the so-called equilibrium ‘semi-experimental’ structures. The range of application of the “top-down” approach can be further extended by employing either the *method of predicate observations*¹⁶ or the *template approach*,¹⁵ in which QC computations are also used to complete the set of experimental data. Whenever the lack of experimental information is too extended, one can resort to the “bottom-up” approach, which consists in verifying the computed equilibrium geometry by means of a comparison between calculated and experimental spectroscopic parameters that probe structural characteristics.

In the following we offer a feeling of the status and perspectives of the quest for reliable structures of molecular systems and their complexes by means of some case studies taken from the work of the authors and their groups.

The joint theory-spectroscopy strategy

In this section the essential details of the joint theory-spectroscopy approach are presented for both the “top-down” and “bottom-up” strategies, and the relevant literature addressed.

The semi-experimental approach

Rotational spectroscopy and its leading terms, i.e., the rotational constants, are at the basis of the semi-experimental approach. Indeed, a generic rotational constant is inversely proportional to the

corresponding element of the inertia tensor \mathbf{I} (in atomic units):¹⁷

$$B^i = \frac{1}{2I_i}, \quad (1)$$

with I_i only depending on molecular geometry and isotopic masses, and i denoting the inertial axis. However, structural determinations from the knowledge of rotational constants are hampered by the limited number of data with respect to the number of geometrical parameters and by the proper account of vibrational effects. The first limitation can be overcome by considering for a given molecule different isotopic species in order to increase the number of available data.¹⁷ However, rotational constants obtained from experiment refer to the vibrational ground state and the structures derived directly from them are vibrationally averaged structures (like r_0 and r_s),¹⁷ therefore depending on the isotopic species considered. On the contrary, the equilibrium geometries exclude vibrational effects in a rigorous manner and are independent of the considered isotopic species. The so-called semi-experimental (SE) equilibrium geometry (r_e^{SE}) is obtained by a least-squares fit of experimental vibrational ground state rotational constants (B_0) of different isotopologues corrected by computed vibrational (ΔB_{vib}) and (often neglected) electronic (ΔB_{el}) contributions:

$$B_e^i = B_0^i - \Delta B_{vib}^i - \Delta B_{el}^i = B_0^i + \frac{1}{2} \sum_n \alpha_n^i - \frac{m_e}{m_p} g^i B_e^i \quad (2)$$

In the above equation, α_n^i denotes a vibration-rotation interaction constant (n is the normal mode),¹⁸ m_e and m_p are the mass of the electron and proton, respectively, and g^i denotes the diagonal element of the rotational g -factor along the i inertial axis.

Rotational spectroscopy experiments need to be carried out in the gas phase. However, in the last decades limitations related to solid compounds have been overcome by employing laser ablation techniques, which allow them to be vaporized without being decomposed. This technique thus paves the way toward the determination of the molecular structure for biomolecules and metal-containing compounds (see, for example, Refs.^{19–24}).

If the number of available isotopologues of a molecular system is not sufficient to allow a robust evaluation of all structural parameters, the lack of information can affect both the quality of the fit and the accuracy and reliability of the parameters determined. The simplest solution of this problem is to fix a subset of parameters to values obtained by means of accurate QC calculations.

A more flexible approach, the so-called *method of predicate observations*,^{2,16,25,26} can be indeed used, which is based on the use of estimated structural parameters as additional input data in the fitting procedure. These estimates are the so-called predicate observations and are weighted according to the accuracy they are expected to meet. For a detailed account, the reader is referred to Ref.¹⁶ A key issue is thus the evaluation of accurate structural parameters to be either kept fixed in the fitting procedure or used as predicates. The so-called *template approach*¹⁵ has been introduced to further extend the size of molecular systems amenable to highly accurate molecular structure determinations. If for a similar molecule or a smaller molecule identical to one of the system moieties an accurate equilibrium geometry can be obtained either experimentally or computationally, then this can be used as template-molecule (TM) for deriving the missing geometrical parameters:

$$r_e(\text{fixed/predicate}) = r_e + \Delta\text{TM} \quad (3)$$

where ΔTM is defined as

$$\Delta\text{TM} = r_e^{\text{accu}}(\text{TM}) - r_e(\text{TM}) \quad (4)$$

r_e is the geometrical parameter of interest calculated at the same level of theory for both the molecule under consideration and that chosen as reference. The level of theory obviously depends on the size of the molecule to be characterized, but usually relies on density functional theory (DFT). Alternatively, linear regressions between computed and experimental bond lengths have been determined and validated.^{27,28}

The “bottom-up” strategy

The idea at the basis of the so-called *bottom-up* strategy is to obtain a validation of the accuracy of the equilibrium geometry computationally determined from the analysis of different spectroscopic parameters that depend on the molecular structure. As mentioned above, rotational constants are inversely proportional to the inertia tensor, which contains the structural information, i.e. the mass distribution in the molecule. Another set of parameters of particular relevance is provided by the nuclear quadrupole coupling constants, χ_{ii} ,¹⁷ which strongly depend on the intramolecular interactions, thus permitting, for example, to distinguish among conformers and/or isomers with

similar mass distributions, i.e. similar rotational constants.²⁴

In general terms, if a very good agreement is observed between the computed and experimental values of those spectroscopic parameters that strongly depend on the molecular structural properties, then an indirect confirmation of the reliability and accuracy of the computed equilibrium geometry is obtained.^{11–13} As already noted for the semi-experimental approach, the proper account of vibrational effects is critical.

Ab initio structures: the composite approaches

Joint experimental-computational approaches are often validated or complemented by accurate ab initio results. To fulfil the accuracy requirements, composite approaches are usually employed because they allow for accounting for both electron correlation and basis-set effects at the best possible level. Indeed, relying on the additivity approximation, the various contributions are evaluated separately at the highest possible level and then combined together. The most physically sound is the *gradient* scheme, with additivity directly applied to energy gradients. Within this family, the CCSD(T)/CBS+CV considers the extrapolation to the CBS limit and CV corrections at the CCSD(T)²⁹ level,^{7,8} s affordable for medium-sized molecules, like glycine⁹ and pyruvic acid,³⁰ and is well recognized to provide bond distances with an accuracy of 0.001-0.002 Å and angles accurate to 0.05-0.1 degrees (see, for example, Refs.^{6–9,30,31}).

Recently, the so-called “cheap” *geometry* scheme has been set up in order to increase the size of systems amenable to accurate structure determinations by reducing their computational cost.^{9,11–13} This very effective approach starts from a geometry optimization at the CCSD(T)/cc-pVTZ level within the frozen-core approximation and adds CBS, CV and aug (i.e. effect of diffuse functions) contributions by means of independent geometry optimizations using second-order Møller-Plesset theory (MP2)³² and basis sets of triple- and quadruple-zeta quality. Several benchmark studies have shown that the accuracy of the “cheap” approach³³ is close to that of the CCSD(T)/CBS+CV reference.^{9,30} In passing, we note that, in addition to those briefly described above, there are several other composite schemes available in the literature, among which the so called focal point approach³⁴ is perhaps the most widely used.

The last comment concerns the quantities to be computed within the different approaches (e.g.

cubic force constants for vibrational corrections, electronic g-tensor, electric field gradients for nuclear quadrupole coupling constants). It should be noted that they are available at different levels of theory within standard calculations and QC program packages. The references in this account are Gaussian³⁵ and CFOUR³⁶ for DFT and CCSD(T) calculations, respectively.

From small systems to building-blocks of biomolecules

In this section, after a short discussion on the accuracy of the semi-experimental approach illustrated by some representative examples, we proceed to address the construction of a quite large database of accurate structures, and then to face the problem of extending the dimension of the systems amenable to accurate structural determinations by means of the template approach and the method of predicate observables.

The semi-experimental approach: from the single result to a database

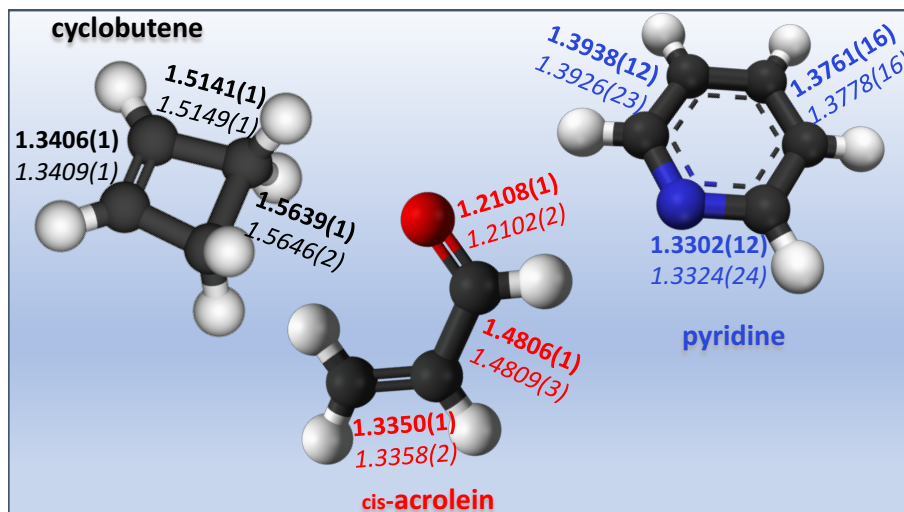


Figure 1: Selection of r_e^{SE} bond lengths of cyclobutene, cis-acrolein, and pyridine obtained using CCSD(T)/cc-pVTZ (in boldface) and B3LYP/SNSD (in italics) vibrational corrections.¹⁵

The reliability of vibrational contributions obtained from QC computations has been demonstrated in the pioneering work by Pawłowski et al.:³⁷ as a matter of fact an accuracy of 0.001 Å on bond lengths can be obtained as soon as a correlated ab initio method in conjunction with

triple-zeta quality basis sets is used. Several subsequent studies have shown that methods rooted in the DFT in combination with double-zeta quality basis sets are able to reach the same accuracy at a greatly reduced computational cost, thus extending the applicability of this approach to large systems.^{15,27} In particular, in Ref.,¹⁵ a systematic benchmark study on small molecules has been carried out to evaluate the performance of the B3LYP/SNSD level of theory in the derivation of vibrational corrections. This evaluation was based on the comparison with vibrational corrections computed at the CCSD(T) level in conjunction with basis sets of at least triple-zeta quality. The overall conclusion was that B3LYP/SNSD corrections lead to results that reproduce very well the best r_e^{SE} geometries. As an example, Figure 1 collects the results for cyclobutene, the cis form of acrolein, and pyridine, thus comparing a selection of r_e^{SE} bond lengths determined using CCSD(T) and B3LYP vibrational corrections. It is observed that the bond lengths agree within 0.001 Å, with the exception of a very few specific cases. For all examples of Figure 1, the electronic corrections to rotational constants were also considered,¹⁵ which are for example mandatory to fulfil the planarity condition for the moments of inertia¹⁷ in cis-acrolein.

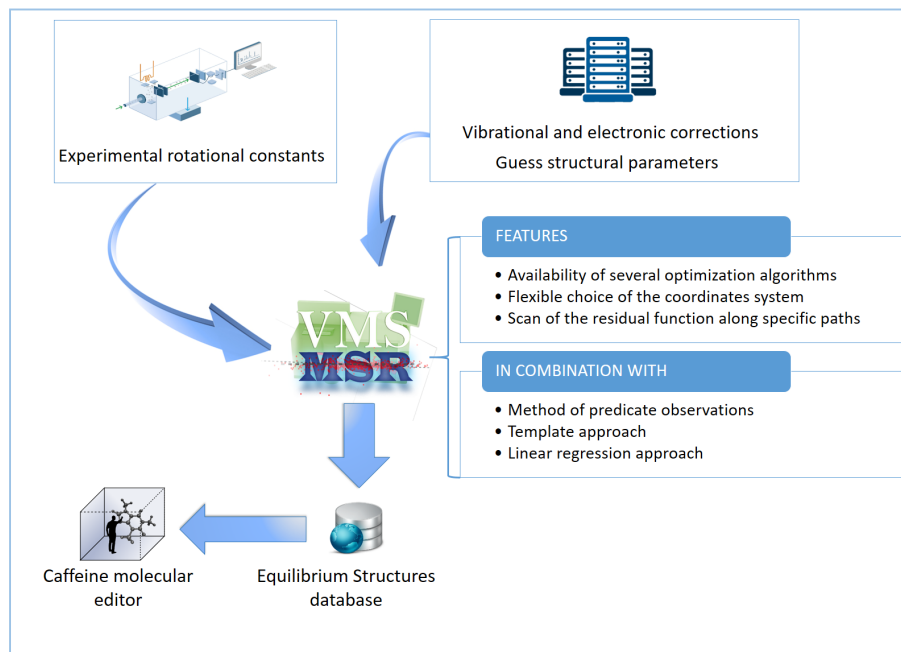


Figure 2: Block diagram of the VMS-MSR software.

A key point in the determination of SE equilibrium structures is the robustness of the fitting program. Furthermore, an important step toward a widespread use of the semi-experimental ap-

proach is the availability of an user-friendly tool. Both requirements are well fulfilled by the recent “Molecular Structure Refinement” (MSR) software,¹⁶ which is specifically designed for computing equilibrium structures based on a nonlinear least-squares fit of the moments of inertia for a given set of isotopologues. Among the wide range of features of the MSR software, an important one is the implementation of a black-box generation of a set of totally symmetric vibrational modes (A_1 internal coordinates) to be used in the structural determination. Indeed, as pointed out by Demaison et al.,² the most effective choice to describe the molecular configuration is by means of the set of A_1 coordinates selected from a set of non-redundant internal symmetry coordinates. In addition, the *method of predicate observations* has also been implemented in the MSR program. While the reader is referred to Ref.¹⁶ for all details, Figure 2 shows the program structure of MSR, which is part of the virtual multifrequency spectrometer (VMS).³⁸

As a sort of conclusion, Ref.¹⁵ demonstrated that the computational bottleneck of the semi-experimental protocol, which is the calculation of the cubic force field required for vibrational corrections, could be overcome by using the B3LYP/SNSD level. Indeed, the differences in the r_e^{SE} geometrical parameters obtained using B3LYP/SNSD and CCSD(T) vibrational corrections show a nearly Gaussian distribution with standard deviation and mean absolute error of 0.0009 and 0.0007 Å for bond distances and of 0.08 and 0.05 degrees for angles, respectively.¹⁵ The small standard deviations shown by B3LYP/SNSD vibrational corrections are therefore fully satisfactory for accurate geometrical parameter determinations, the values given above being clearly comparable with the intrinsic errors of the fitting procedure itself. The evaluation of r_e^{SE} for a set of 47 molecules containing the most common organic moieties led to the definition of a database of accurate molecular geometries,^{15,27} the so-called B3se database, to be used as references for benchmark calculations as well as for the validation of low-level models affordable for large systems. In particular, Refs.^{39,40} made use of the B3se database for benchmarking recently developed density functionals. In subsequent studies, it has been shown that even better results can be obtained employing the B2PLYP functional in conjunction with a triple-zeta basis set²⁷ (B2se set) and the database was extended to compounds containing sulfur.²⁸ The reader is referred to the database website (<http://smart.sns.it/molecules/>), where all information can be downloaded.

Even if the topic will not be further addressed, it is worthwhile mentioning that semi-experimental equilibrium constants can also be used for benchmarking low level, in particular DFT, models. A

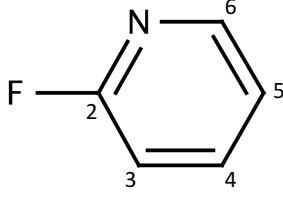
significant example in this respect is provided by Ref.⁴¹ The last comment concerns other experimental techniques for structural determinations, which also suffer from the proper analysis of vibrational effects. Among them, gas-phase electron diffraction and x-ray crystallography deserve a particular mention, the latter also showing the additional issue of crystal packing to be taken into account. In these cases, vibrational corrections can be evaluated to correct computed equilibrium geometries in order to derive structures that are directly comparable with experimental values.⁴²

Extension to building-blocks of biomolecules

A great challenge of computational chemistry is the extension of the accuracy obtainable with the semi-experimental approach to larger systems, like building-blocks of biomolecules. As already mentioned, the key point of the semi-experimental approach is to have a sufficient number of experimental data in order to obtain a balanced fit, which means not only to have a number of rotational constants larger than the number of molecular parameters, but also to have data for isotopic substitutions involving all nuclei. This becomes clearly not possible when the molecular size and topological complexity increase because of the large number of isotopologues then required. The *template approach* introduced above provides a solution when a pure, accurate ab initio determination cannot be carried out. An illustrative example is provided by 2-fluoropyridine. The first determination of the semi-experimental equilibrium structure of 2-fluoropyridine is reported in Ref.¹⁵. Because of the limited number of available experimental data, not all structural parameters could be determined. Instead, some parameters were fixed using the template approach, with pyridine as template molecule. In Table 1 two fits from Ref.¹⁵ are given: in the first one the CCSD(T)/CBS+CV level has been considered for evaluating r_e of equations (3) and (4), while in the second fit r_e was calculated at the B3LYP/SNSD level. The third fit collects the results when the *template approach* is not employed and the non-determinable parameters are kept fixed at their B3LYP/SNSD values. The comparison of the third fit to the first two clearly points out that fixing parameters to not sufficiently accurate values strongly affects the outcome of the fit itself, thus leading to inaccurate results. However, the comparison between the first two fits demonstrates that the B3LYP/SNSD level is suitable once employed in conjunction with the *template approach*.

Another strategy to overcome the lack of information and possible strong correlations between

Table 1: Partial r_e^{SE} of 2-fluoropyridine (bond lengths in Å, angles in degrees).

			
Parameter	Ref. ¹⁵		This work
	Fit 1 ^a	Fit 2 ^b	Fit 3 ^c
$r(\text{N-C2})$	1.3179(38)	1.3179(38)	1.3057(59)
$r(\text{N-C6})$	1.3402(5)	1.3402(5)	1.3419(6)
$r(\text{C2-C3})$	1.3796(40)	1.3796(39)	1.3989(60)
$r(\text{C4-C5})$	1.3961(11)	1.3961(11)	1.3921(17)
$r(\text{C5-C6})$	1.3826(10)	1.3826(10)	1.3859(15)
$r(\text{C2-F})$	1.3356(3)	1.3356(3)	1.3334(4)
$r(\text{C3-H})$	1.0781 ^d	1.0755 ^d	1.0839 ^d
$r(\text{C4-H})$	1.0801 ^d	1.0796 ^d	1.0860 ^d
$r(\text{C5-H})$	1.0788 ^d	1.0783 ^d	1.0848 ^d
$r(\text{C6-H})$	1.0809 ^d	1.0811 ^d	1.0876 ^d
$\angle(\text{C6-N-C2})$	116.15(9)	116.15(9)	116.35(13)
$\angle(\text{N-C2-C3})$	126.18(2)	126.18(2)	126.33(2)
$\angle(\text{C5-C6-N})$	123.47(3)	123.47(3)	123.37(4)
$\angle(\text{C4-C5-C6})$	118.36(1)	118.36(1)	118.38(1)
$\angle(\text{C3-C2-F})$	118.78(33)	118.77(32)	117.67(51)
$\angle(\text{C2-C3-H})$	120.38 ^d	120.48 ^d	120.60 ^d
$\angle(\text{C3-C4-H})$	120.62(26)	120.64(26)	121.33(28)
$\angle(\text{C6-C5-H})$	120.26 ^d	120.25 ^d	120.37 ^d
$\angle(\text{C5-C6-H})$	120.86 ^d	120.86 ^d	120.97 ^d

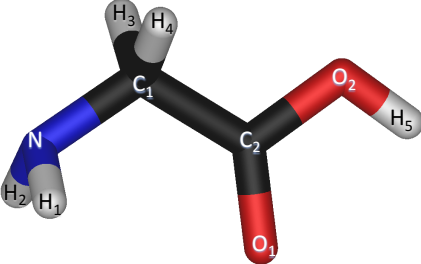
^a $r_e(\text{fixed}) = r_e(\text{CCSD(T)/CBS+CV}) + \Delta\text{TM}$; $\Delta\text{TM} = r_e^{SE}(\text{B3LYP/SNSD}) - r_e(\text{CCSD(T)/CBS+CV})$

^b $r_e(\text{fixed}) = r_e(\text{B3LYP/SNSD}) + \Delta\text{TM}$; $\Delta\text{TM} = r_e^{SE}(\text{B3LYP/SNSD}) - r_e(\text{B3LYP/SNSD})$

^c $r_e(\text{fixed}) = r_e(\text{B3LYP/SNSD})$

^d Fixed parameters.

Table 2: Semi-experimental equilibrium structure of glycine Ip (bond lengths in Å, angles in degrees).



Parameter	Ref. ⁹	Ref. ¹⁶		
	Fit 1 ^a	Fit 2 ^b	Fit 3 ^c	Fit 4 ^d
$r(\text{C2-O2})$	1.348(3)	1.349(1)	1.350(3)	1.347(3)
$r(\text{C2-O1})$	1.203(5)	1.203(1)	1.203(2)	1.203(3)
$r(\text{O2-H5})$	0.9645fix	0.965(2)	0.968(7)	0.97(2)
$r(\text{C1-C2})$	1.513(4)	1.5133(7)	1.513(1)	1.513(1)
$r(\text{C1-N})$	1.443(1)	1.4429(9)	1.442(2)	1.442(3)
$r(\text{C1-H3})$	1.0908(6)	1.0907(3)	1.0907(3)	1.0907(3)
$r(\text{N-H1})$	1.0104(9)	1.011(2)	1.011(7)	1.02(1)
$\angle(\text{O1-C2-O2})$	123.00(3)	123.03(6)	123.1(1)	122.3(2)
$\angle(\text{C1-C2-O2})$	111.5(4)	111.44(8)	111.5(2)	111.5(2)
$\angle(\text{N-C1-C2})$	115.3(1)	115.28(5)	115.3(1)	115.3(2)
$\angle(\text{H3-C1-C2})$	[107.36]	107.39(4)	107.38(5)	107.37(5)
$\angle(\text{H1-N-C1})$	[110.10]	109.9(2)	109.7(6)	110.2(8)
$\angle(\text{H5-O2-C2})$	106.64fix	106.7(2)	106.6(7)	107.1(8)
$\delta(\text{H4-C1-C2-O1})$	[123.21]	123.20(4)	123.21(4)	123.21(4)
$\delta(\text{H2-N-C1-C2})$	[57.43]	57.4(2)	57.2(6)	57.(1)

^a Partial r_e^{SE} with fixed parameters at the CCSD(T)/CBS+CV level. For details, see Ref. ⁹

^b Fit performed using CCSD(T)/CBS+CV predicate observations. The uncertainties assigned to bond distances, bond angles, and dihedral angles are: 0.001 Å, 0.1 deg., and 1.0 deg., respectively. For details, see Ref. ¹⁶

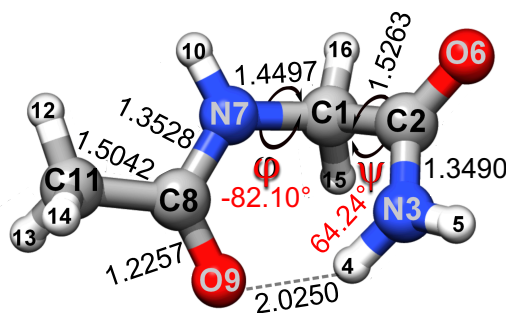
^c Fit performed using B2PLYP/cc-pVTZ predicate observations. The uncertainties assigned to bond distances, bond angles, and dihedral angles are: 0.005 Å, 0.5 deg., and 2.0 deg., respectively. For details, see Ref. ¹⁶

^d Fit performed using B3LYP/SNSD predicate observations. The uncertainties assigned to bond distances, bond angles, and dihedral angles are: 0.01 Å, 0.5 deg., and 2.0 deg., respectively. For details, see Ref. ¹⁶

parameters, which can greatly affect the quality of the fit, is the so-called *method of predicate observations* introduced above. In this approach the fitting procedure remains mostly unchanged, except for the fact that the inclusion of predicates leads to an augmented set of data for the fitting itself. To illustrate this method, the equilibrium structure determination of the most stable conformer of glycine (Ip), which has been the object of several studies (see Refs.^{9,16} and references therein), has been considered. The number of available isotopic species being not sufficient to characterize all internal degrees of freedom, a partial semi-experimental equilibrium geometry was determined in Ref.,⁹ with the non-determinable structural parameters kept fixed at high-level computed values. In Ref.,¹⁶ the so-called *method of predicate observations* was employed with the set of semi-experimental moments of inertia used in Ref.⁹ augmented by the computed data mentioned above used as predicate observations. The r_e^{SE} structure has then been obtained by means of a weighted least-squares calculation, in which the weights have been defined on the basis of experimental and assigned uncertainties for the semi-experimental moments of inertia and the predicate observations, respectively. In detail, the computed values were obtained at the so-called CCSD(T)/CBS+CV level of theory. The two geometries are compared in Table 2 (first two fits). However, these two fits have been carried out by employing structural values at a high level of theory, which is not affordable for larger systems. For this reason, in Ref.¹⁶ the *method of predicate observations* has also been applied by using predicate observations at both the B2PLYP/cc-pVTZ and B3LYP/SNSD levels. The results are collected in the third and fourth columns of Table 2. It is first of all noted that all fits provide very similar results. This is particularly interesting for the three fits obtained with the *method of predicate observations* because of the huge difference in the computational cost between the CCSD(T)/CBS+CV level of theory on one side and the B2PLYP/cc-pVTZ and B3LYP/SNSD levels on the other. This example thus demonstrates that the extremely cheap B3LYP/SNSD level of theory can be used to provide reliable predicate observations able in turn to lead to accurate semi-experimental equilibrium structures.

Whenever the lack of data is even more significant, the limited experimental data can be used to obtain confirmation of the accuracy of the computed equilibrium structure in the framework of the so-called *bottom-up* approach. To illustrate this strategy, the investigation of the rotational spectrum of N-acetyl-glycinamide (Ac-Gly-NH₂), also denoted as glycine dipeptide analogue, is briefly discussed. For N-acetyl-glycinamide, two conformers are found to be stable in the gas

Table 3: Experimental and calculated spectroscopic parameters (in MHz) of the C_7 N-acetylglycinamide conformer.^a



Parameter	Experiment	Calculated ^b
A_0	4469.1107(11)	4480.5
B_0	1214.23622(27)	1210.6
C_0	1081.26289(20)	1082.9
$\chi_{aa}(\text{N7})$	1.2553(38)	1.258
$\chi_{bb}(\text{N7})$	0.717(74)	0.752
$\chi_{cc}(\text{N7})$	-1.972(74)	-2.008
$\chi_{aa}(\text{N3})$	0.3887(69)	0.432
$\chi_{bb}(\text{N3})$	1.593(10)	1.603
$\chi_{cc}(\text{N3})$	-1.981(10)	-2.035

^a Ref.¹³

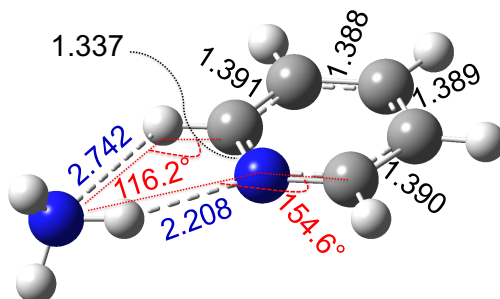
^b Best-estimated equilibrium rotational and nitrogen quadrupole coupling constants (from the “cheap” geometry scheme) augmented by vibrational corrections at the B3LYP/SNSD level. In the figure: a selection of structural parameters (distances in Å, angles in degrees).

phase: C_5 and C_7 , with the latter being the most stable by about 200 cm^{-1} . In Ref.,¹³ the “cheap” geometry scheme was applied to the spectroscopic characterization aiming at assigning the rotational spectrum. For the C_7 conformer, a selection of structural parameters is given in the figure of Table 3, which summarizes the comparison between experiment and theory for the rotational and nitrogen quadrupole coupling constants. A very good agreement is noted, with discrepancies of the order of 0.3% for the former and 1-3% for the latter. This very good agreement clearly indicates that the computational scheme employed was able to describe well both the molecular structure and the intramolecular interactions, thus suggesting a very good accuracy for the structural parameters, also for the backbone torsion angles φ and ψ , which describe the conformation of the C_7 and C_5 rotamers. Indeed, since both the amine and amide functional groups are involved in the φ and ψ angles, the nitrogen quadrupole coupling constants change by varying the dihedral angle values.

A step toward the condensed phase

In the last decades, a great effort has been devoted to understand the nature of weak and/or electrostatic interactions in condensed phase. The *bottom-up* strategy applied above to the elucidation of the structure of a glycine dipeptide analogue can also be successfully employed to the accurate description of intermolecular interactions. These range from those formed by the system under consideration with one or more molecules of water or other solvents to those established between two or more molecular systems themselves. An example for the first case is provided by the joint experimental-computational investigation of the pyridine-ammonia complex.⁴³ For the first time, the “cheap” geometry scheme was applied to a weakly bonded molecular complex, and its accuracy and reliability was tested by means of a high-resolution rotational spectroscopy study. Key results are collected in Table 4: they demonstrate that the “cheap” approach is suitable for studying molecular complexes and is able to lead to accurate structural determinations. As above, relying on a *bottom-up* strategy, the good agreement observed for rotational and nitrogen quadrupole coupling constants together with the specialistic literature on this topic^{9,11,12} suggest that the structural parameters, also those involved in the intermolecular interaction, have an estimated accuracy of about 0.001-0.002 Å for bond lengths and around 0.1-0.2 degrees for angles.

Table 4: Experimental and calculated spectroscopic parameters (in MHz) of the pyridine-ammonia complex.^a



Parameter	Experiment	Calculated ^b
A_0	5872.300(1)	5875.20
B_0	1387.6497(2)	1386.28
C_0	1128.7913(1)	1127.93
N (Pyridine)		
χ_{aa}	3.747(5)	3.92
$\chi_{bb} - \chi_{cc}$	2.860(8)	2.86
N (Ammonia)		
χ_{aa}	1.437(5)	1.42
$\chi_{bb} - \chi_{cc}$	4.792(8)	4.49

^a Ref.⁴³

^b Best-estimated equilibrium rotational and nitrogen quadrupole coupling constants (from the “cheap” geometry scheme) augmented by vibrational corrections at the B2PLYP-D3/m-aug-cc-pVTZ-dH level. In the figure: a selection of structural parameters (distances in Å, angles in degrees).

As mentioned above, recent experimental techniques and in particular the employment of laser ablation in conjunction with broadband microwave spectroscopy, allow the investigation of systems of increasing dimension and complexity, also involving intermolecular interactions. This is testified by several works reported in the literature (see, for example, Refs.^{44–47}). More important in the context of this account is that the sensitivity of these new techniques permits the observation of isotopically substituted species in natural abundance, thus allowing the determination of effective structures, like r_0 and r_s . This is, for example, demonstrated in Ref.⁴⁷ for the complexes formed by β -propiolactone and one to five water molecules and in Ref.⁴⁵ for the phenol dimer and trimer. However, as already pointed out, effective structures are not sufficiently accurate and in some cases they also fail in providing qualitative results (see, for example, Ref.⁶). Furthermore, as shown in Refs.,^{45,47} when the systems become large, very often only partial structures can be obtained, with non-determinable parameters fixed at computed equilibrium values. The semi-experimental approach combined with either the *template approach* and/or the *method of predicate observations* can solve inconsistency issues in the structural determination and allows for obtaining accurate and quantitative descriptions of intermolecular bonds. However, methodological developments and benchmark studies are still in progress in our groups because anharmonic force field computations for flexible systems, like molecular complexes, are not at all straightforward.

Perspectives

Despite the great advances and accuracy in structural determinations seen in the last decades, several challenging cases are still present. In addition to the determination of semi-experimental equilibrium structures for molecular complexes, contemporary challenges for a joint theory-spectroscopy approach are:

- *Correct treatment of large amplitude motions.* The presence of large amplitude motions poses challenges for both theory and experiment. The former has to face the typical inaccuracy in the determination of dihedral angles, the latter usually lacks of the appropriate Hamiltonian for the analysis of the spectroscopic data.

- *Heavy metal(s) in the system under investigation.* While an effective treatment of relativistic effects in QC computations can be considered an accomplished task, heavy metal-containing compounds are often investigated in a condensed phase. This poses the challenge of a correct model to interpret the spectroscopic results. An illustrative example is provided by cisplatin,⁴⁸ for which the vibrational spectroscopic features can be correctly interpreted only if the essential intermolecular interactions are accounted for.

While a thorough discussion on the challenges above is beyond the scope of the present account, it is noted that methodologies to properly address them are under development in our research groups.

Biographical Information

Cristina Puzzarini is Full Professor of Physical Chemistry at the University of Bologna. After a Ph.D. in Theoretical Chemistry (1997), in 2000 she moved back to the Millimeter-/Submillimeter-wave Spectroscopy lab, where in 1993 she got her MSc degree in Chemistry. Her research lines cover both computational and experimental aspects in the field of High-Resolution Molecular Spectroscopy.

Vincenzo Barone is Full Professor of Theoretical and Computational Chemistry at the Scuola Normale Superiore in Pisa (from 1994 to 2007 at the Federico II University of Naples), where he is also the Director. He is Fellow of the Royal Society of Chemistry, Accademia dei Lincei and IAQMS and has been President of the Italian Chemical Society (SCI). His main research interests include DFT, solvation and computational spectroscopy.

Acknowledgments

This work was supported by Italian MIUR (PRIN 2015) and by the University of Bologna (RFO funds). The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. [320951]

References

- (1) Domenicano, A., Hargittai, I., Eds. *Accurate Molecular Structures: Their Determination and Importance*; 1992.
- (2) Demaison, J., Boggs, J. E., Császár, A. G., Eds. *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; CRC Press, Taylor & Francis Group, 2011.
- (3) Barone, V., Ed. *Computational Strategies for Spectroscopy, from Small Molecules to Nano Systems*; John Wiley & Sons, Inc. Hoboken: New Jersey, 2011.
- (4) Puzzarini, C.; Biczysko, M. In *Structure Elucidation in Organic Chemistry*; Cid, M. M., Bravo, J., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA, 2015; pp 27–64.
- (5) Puzzarini, C. Rotational spectroscopy meets theory. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6595–6607.
- (6) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273–367.
- (7) Heckert, M.; Kállay, M.; Gauss, J. Molecular Equilibrium Geometries Based on Coupled-Cluster Calculations Including Quadruple Excitations. *Mol. Phys.* **2005**, *103*, 2109–2115.
- (8) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. Basis-Set Extrapolation Techniques for the Accurate Calculation of Molecular Equilibrium Geometries Using Coupled-Cluster Theory. *J. Chem. Phys.* **2006**, *125*, 044108/1–10.
- (9) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. The Performance of Composite Schemes and Hybrid CC/DFT Model in Predicting Structure, Thermodynamic and Spectroscopic Parameters: the Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094–10111.
- (10) Puzzarini, C. Accurate molecular structures of small- and medium-sized molecules. *Int. J. Quantum Chem.* **2016**, *116*, 1513–1519.

- (11) Puzzarini, C.; Barone, V. Extending the Molecular Size in Accurate Quantum-Chemical Calculations: the Equilibrium Structure and Spectroscopic Properties of Uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189–7197.
- (12) Puzzarini, C.; Biczysko, M.; Barone, V.; Pena, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic properties of nucleobases: a combined computational-microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965–16975.
- (13) Puzzarini, C.; Biczysko, M.; Barone, V.; Largo, L.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Characterization of the Peptide Linkage in the Gas Phase: a Joint Quantum-chemical and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* **2014**, *5*, 534–540.
- (14) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic Force Constants and Equilibrium Geometry of Methane from Hartree–Fock and Correlated Wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077–5085.
- (15) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules. *J. Phys. Chem. A* **2015**, *119*, 2058–2082.
- (16) Mendolicchio, M.; Penocchio, E.; Licari, D.; Tasinato, N.; Barone, V. Development and Implementation of Advanced Fitting Methods for the Calculation of Accurate Molecular Structures. *J. Chem. Theo. Comp.* **2017**, *13*, 3060–3075.
- (17) Gordy, W.; Cook, R. L. In *Microwave Molecular Spectra, 3rd Edition*; Weissberger, A., Ed.; Wiley: New York, 1984.
- (18) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Eds.; Academic: New York, 1972; Chapter Vibration-Rotation Structure in Asymmetric- and Symmetric-Top Molecules, p 115.

- (19) Lesarri, A.; Mata, S.; Lopez, J. C.; Alonso, J. L. A laser-ablation molecular-beam Fourier-transform microwave spectrometer: The rotational spectrum of organic solids. *Rev. Sci. Instrum.* **2003**, *74*, 4799–4804.
- (20) Alonso, J. L.; López, J. C. Microwave spectroscopy of biomolecular building blocks. *Top. Curr. Chem.* **2015**, *364*, 335–401.
- (21) Banser, D.; Schnell, M.; Grabow, J.-U.; Cocinero, E. J.; Lesarri, A.; Alonso, J. L. The Internuclear Potential, Electronic Structure, and Chemical Bond of Tellurium Selenide. *Angew. Chem. Int. Ed.* **2005**, *44*, 6311–6315.
- (22) Brücken, S.; Müller, H. S. P.; Menten, K. M.; McCarthy, M. C.; Thaddeus, P. The Rotational Spectrum of TiO₂. *Astrophys. J.* **2008**, *676*, 1367–1371.
- (23) Mück, L. A.; Thorwirth, S.; Gauss, J. The semi-experimental equilibrium structures of Al-CCH and AlNC. *Journal of Molecular Spectroscopy* **2015**, *311*, 49–53.
- (24) Peña, I.; Sanz, M. E.; Lopez, J. C.; Alonso, J. L. Preferred Conformers of Proteinogenic Glutamic Acid. *J. Am. Chem. Soc.* **2012**, *134*, 2305–2312.
- (25) Demaison, J. Experimental, semi-experimental and ab initio equilibrium structures. *Mol. Phys.* **2007**, *105*, 3109–3138.
- (26) Vogt, N.; Demaison, J.; Ksenafontov, D. N.; Rudolph, H. D. A Benchmark Study of Molecular Structure by Experimental and Theoretical Methods: Equilibrium Structure of Thymine from Microwave Rotational Constants and Coupled-cluster Computations. *J. Mol. Struct.* **2014**, *1076*, 483–489.
- (27) Penocchio, E.; Piccardo, M.; Barone, V. Semiexperimental Equilibrium Structures for Building Blocks of Organic and Biological Molecules: The B2PLYP Route. *J. Chem. Theo. Comp.* **2015**, *11*, 4689–4707.
- (28) Penocchio, E.; Mendolicchio, M.; Tasinato, N.; Barone, V. Structural features of the carbon-sulfur chemical bond: a semi-experimental perspective. *Can. J. Chem.* **2016**, *94*, 1065–1076.

- (29) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (30) Barone, V.; Biczysko, M.; Bloino, J.; Cimino, P.; Penocchio, E.; Puzzarini, C. CC/DFT Route toward Accurate Structures and Spectroscopic Features for Observed and Elusive Conformers of Flexible Molecules: Pyruvic Acid as a Case Study. *J. Chem. Theory Comput.* **2015**, *11*, 4342–4363.
- (31) Barone, V.; Biczysko, M.; Puzzarini, C. Quantum Chemistry Meets Spectroscopy for Astrochemistry: Increasing Complexity toward Prebiotic Molecules. *Acc. Chem. Res.* **2015**, *48*, 1413–1422.
- (32) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (33) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate Molecular Structures and Infrared Spectra of trans-2,3-dideuteriooxirane, Methyloxirane, and trans-2,3-dimethyloxirane. *J. Chem. Phys.* **2014**, *141*, 034107/1–17.
- (34) Allen, W. D.; East, A. L. L.; Csaászai, A. G. In *Structures and Conformations of Non-Rigid Molecules*; Laane, J., Dakkouri, M., van der Veken, B., Oberhammer, H., Eds.; Kluwer: Dordrecht, 1993; p 343.
- (35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.;

- Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Revision A.03. 2016; Gaussian Inc. Wallingford CT.
- (36) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. CFOUR A quantum chemical program package. 2011; with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, F. Lipparini, D. Matthews, T. Metzroth, L. A. Mueck, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almloef and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wuelen. For the current version, see <http://www.cfour.de>.
- (37) Pawłowski, F.; Jørgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. Molecular Equilibrium Structures from Experimental Rotational Constants and Calculated Vibration–Rotation Interaction Constants. *J. Chem. Phys.* **2002**, *116*, 6482–6496.
- (38) Barone, V. The virtual multifrequency spectrometer: a new paradigm for spectroscopy. *WIREs Comput. Mol. Sci.* **2016**, *6*, 86–110.
- (39) Bremond, E.; Savarese, M.; Su, N. Q.; Perez-Jimenez, A. J.; Xu, X.; Sancho-Garcia, J. C.; Adamo, C. Benchmarking Density Functionals on Structural Parameters of Small-/Medium-Sized Organic Molecules. *J. Chem. Theo. Comp.* **2016**, *12*, 459–465.
- (40) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn-Sham global-hybrid exchange-correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.* **2016**, *7*, 5032–5051.
- (41) Risthaus, T.; Steinmetz, M.; Grimme, S. Implementation of nuclear gradients of range-separated hybrid density functionals and benchmarking on rotational constants for organic molecules. *J. Comp. Chem.* **2014**, *35*, 1509–1516.

- (42) Waller, M. P.; Bühl, M. Vibrational corrections to geometries of transition metal complexes from density functional theory. *J. Comp. Chem.* **2007**, *28*, 1531–1537.
- (43) Spada, L.; Tasinato, N.; Vazart, F.; Barone, V.; Caminati, W.; Puzzarini, C. Noncovalent Interactions and Internal Dynamics in Pyridine-Ammonia: A Combined Quantum-Chemical and Microwave Spectroscopy Study. *Chem. Eur. J.* **2017**, *23*, 4876–4883.
- (44) Peña, I.; Cabezas, C.; Alonso, J. L. Unveiling epimerization effects: a rotational study of α -d-galactose. *Chem. Commun.* **2015**, *51*, 10115–10118.
- (45) Cocinero, E. J.; Lesarri, A.; Cija, P.; Basterretxea, F. J.; Grabow, J.-U.; Fernández, J. A.; Castaño, F. Ribose Found in the Gas Phase. *Angew. Chem. Int. Ed.* **2012**, *51*, 3119–3124.
- (46) Lesarri, A.; Shipman, S. T.; Neill, J. L.; Brown, G. G.; Suenram, R. D.; Kang, L.; Caminati, W.; Pate, B. H. Interplay of Phenol and Isopropyl Isomerism in Propofol from Broad-band Chirped-Pulse Microwave Spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 13417–13424.
- (47) Perez, C.; Neill, J. L.; Muckle, M. T.; Zaleski, D. P.; Pena, I.; Lopez, J. C.; Alonso, J. L.; Pate, B. H. Water-Water and Water-Solute Interactions in Microsolvated Organic Complexes. *Angew. Chem. Int. Ed.* **2015**, *54*, 979–982.
- (48) Tasinato, N.; Puzzarini, C.; Barone, V. Correct Modeling of Cisplatin: a Paradigmatic Case. *Angew. Chem. Int. Ed.* **2017**, *56*, 13838–13841.